# MSDS 413, Summer 2019, Assignment 5
# Nonstationary Univariate ARMA Models (TS5)

## Introduction

The monthly market liquidity measures are from Professors Pastor and Stambaugh. The data are available from Wharton WRDS and are in the file m-PastorStambaugh.txt.

The following list defines the variables:

- DATE: is the month the data were collected
- PS_LEVEL: levels of aggregate liquidity
- PS_INNOV: innovations in aggregate liquidity
- PS_VWF: traded liquidity factor

The monthly Fama-Bliss bond yields have maturities of 1 and 3 years. The data are available from CRSP and are in the file m-FamaBlissdbndyields.txt.

The following list defines the variables:

- qdate: month of the yields
- yield1: 1 year yields
- yield3: 3-year yields

Your objective is to explore the time series behavior of these data sets including EDA, modeling, model diagnostics, and interpretation.

## Procedure

The following steps are necessary to complete this assignment. Address each and every part and ensure that you cover all the details specified in the questions.

1. **Professors Pastor and Stambaugh** (1.5 points) Consider the monthly market liquidity measure of Professors Pastor and Stambaugh. The data are available from Wharton WRDS and are in the file m-PastorStambaugh.txt. Consider the variable $PS$ level and denote the series by $x_t$.

    1.1. Perform EDA.

    1.2. Build a time series model for $x_t$ (the mean equation) using the model-building process. Write the equation of the model to be fitted (not the fitted model).

    1.3. Identify the largest outlier in the series. Refine the fitted model by using an indicator for the outlier. Write the equation of the refined model.

1.4. Further refine the model by setting the least significant parameter to zero. Write the equation of the revised model to be fitted..

2. **Fama-Bliss bond yields** (1.5 points) Consider the monthly Fama-Bliss bond yields with maturities of 1 and 3 years. The data are available from CRSP and are in the file m-FamaBlissdbndyields.txt. Denote the yields by $y_{1t}$ and $y_{3t}$, respectively.

   2.1. Perform EDA.

   2.2. Build a time series model using the model-building process for the year three ($y_{3t}$) data. For simplicity, you may ignore possible outliers, but describe how you would treat outliers if they were not to be ignored.

   2.3. Fit the following model to the log earnings series:

   **m5 <- arima(xt, order=c(0,1,1), seasonal=list(order=c(0,0,1), period=4))**

   where **xt** denotes the log of the earnings. Write the equation of the fitted model. Compare this model with the model in part 2.2.. Which model is preferred? Why?

   2.4. Use the backtest procedure to compare these same two models via 1-step ahead forecasts. You may use $t = 600$ as the starting forecast origin. Which model is preferred? Why?

3. **Fama-Bliss bond yields** (1.5 points) Consider the monthly Fama-Bliss bond yields with maturities of 1 and 3 years. The data are available from CRSP and are in the file m-FamaBlissdbndyields.txt. Denote the yields by $y_{1t}$ and $y_{3t}$, respectively. The goal is to explore the dependence of the 3-year yield on the 1 year yield.

   3.1. Perform additional EDA if necessary.

   3.2. Fit the linear regression model $y_{3t} = \beta_0 + \beta_1 y_{1t} + e_t$ using the model-building process. Write the equation of the model to be fitted.

   3.3. Repeat by letting $d_{1t} = (1 - B)y_{1t}$ and $d_{2t} = (1 - B)y_{3t}$, where B is the back-shift operator. Here $d_{it},\ i = 1, 2, 3$ denotes the change in monthly bond yields. Consider the linear regression $d_{3t} = \beta d_{1t} + e_t$. Write the equation of the model to be fitted. Refine as needed to achieve an adequate model.

   3.4. Based on the refined model, describe the linear dependence between the bond yields.

4. **Bond yields** (1.5 points) Consider again the bond yields of Part 3. Suppose we are concerned with taking the first difference.

   4.1. Fit an $AR(6)$ model to $y_{3t}$ using $y_{1t}$ as an explanatory variable using the model-building process. Write the equation of the model to be fitted.

   4.2. Refine the model in 4.1. by setting the insignificant coefficients of lags 2 and 5 to zero. Write the equation of the fitted model. Compare this model with that in 4.1..

   4.3. Use the command **polyroot** in R to find the solutions of the characteristic equation of the refined $AR(6)$ model. How many real solutions are there?

4.4. Compute the inverse of the absolute values of the solutions of the characteristic equation. Write the maximum value of the inverses. The maximum should be close to 1, implying that the $AR(6)$ model likely contains a unit root.

5. **Report** (1.5 points) For theFama-Bliss bond yields analyses, describe the expected yields for a predicted period at what confidence.

## Deliverables

See Section Submission Directions below. The assignment deliverables, each in pdf format, are as follows:

- *Only if requested by instructor*
  - The program or script
  - Logs
  - Outputs
- Mandatory
  Data analysis write-up: no programs, logs, or just code outputs.

The data analysis must follow and use the item numbering of each assignment, i.e., use the numbers, say, 1 - 5, with the sub-lettering if used. These deliverables are provided according to the instructions in the Submission Directions section below.

## Submission Directions

### Title Page

Include a title page with your name and the assignment designation. Leave room for instructor comments.

### File Names

The assignment write-up file shall be submitted to Canvas according to the schedule in the syllabus using the item (1) naming convention below. The naming convention is case sensitive. Use letters and numbers as given. **The file name parts have no spaces or other separator charancters**.
TS5Lastname.pdf (submit via Canvas)

The parts are the assignment code, TS5; your lastname with only the first letter capitalized; a period, and lastly, the extension "pdf". Generically,

TS5Lastname.pdf

For example: Suppose your name is Student McStats. Your filename then is:

TS5Mcstats.pdf

The analysis write-up file must be submitted for grading. Each write-up requires a title page for instructor comments. The analysis may use either R or any other statistics package you wish, or if you use more than one package, you must use the germane tables, plots, etc., in a single report. If you use more than one package, differences and similarities should be indicated.

**email: jamie.riggs@northwestern.edu**

Email *ONLY IF REQUESTED* the program (script), log and output as separate pdf files. The R log and output may be combined. The file names shall be as follows:

- The program or script file names
  - TS5LastnameRprog.pdf
- The log file names
  - TS5LastnameRlog.pdf